Semantic Data Modeling

- What is Conceptual Data Modeling
- Entity-Relationship (E-R) Modeling
- Limitations of E-R Modeling
- Object-oriented Modeling: Another semantic model (discussed later in this course)
What Is Conceptual Data Modeling?

A process that represents the entities, relationships, and activities of an enterprise in terms of a set of abstract concepts of a chosen data model for specific purposes.

*Enterprise Modeling, Business Modeling*
Bridge the Gap

Student (ID, Name, Age, Address, GPA)
Instructor (Emp#, Name, Rank, Dept)
Course (Course#, Credits, Title)
Class (Emp#, ID, Course#, Time, Room)
Data Semantics

 Worst Static Information

 - Data -- Entities
 - Associations -- Relationships among entities

 Worst Dynamic Information

 - Activities -- Operations/transactions
 - Integrity constraints -- Business rules/regulations and data meanings
A conceptual data model consists of:

- A collection of formal concepts
- A set of usage rules

Different model has different modeling capability

- **Conventional data modeling**
  - Hierarchical
  - Network
  - Relational

- **Object-Oriented data modeling**
  - E-R
  - EER
  - etc.

- **Semantic data modeling**
  - E-R
  - EER
  - etc.
E-R Modeling

- Introduced by Peter Chen in 1976

- Basic modeling concepts:
  - Entities, entity types, and attributes
  - Relationships
E-R Notation

Entity

Relationship

Attribute

Primary Key
Entities

- An entity is a conceptual object
- Physically exists
- Usually a noun in requirement specification
**Entity Types**

- A collection of similar entities
- An abstraction of "physical" entities
- A noun in requirement specifications
- Having "independent" meaning
Attributes

Properties or characteristics of entities and entity types

- Attribute values -- Properties of entities
- Value set - All acceptable attribute values
- Attributes (definitions) -- Properties of entity types

- A noun or an adjective in requirement specifications
- No "independent" meaning
Key Attributes

One or a group of attributes that can uniquely identify individual entities of an entity type

- A key refers to one or a group of attributes as a whole
- A key attribute is a component attribute of a key
- Key changes with data semantics

An entity type may have several qualified keys

- Primary key -- One of the candidate keys
- Alternate key - Candidate keys not used as the primary key
- Secondary key -- An identifier of records with similar properties of interest

- The primary key attribute(s) is(are) underlined
More Attributes

- **Simple attribute**
  Contains atomic values only

- **Composite attribute**
  Has component attributes

- **Single-valued attribute**
  Has exactly one value per entity

- **Multi-valued attribute**
  Contains repeating values per entity

- **Derived attribute**
  Attribute values computed by means of other attributes
Relationships

Associations among entities

- Relationships -- Associations among entities
- Usually a verb in requirement specification
Relationship Degrees

The number of entity types associated with that relationship.
Relationship Cardinality

How entities are connected through a relationship

- **One-to-One** -- An entity of $E1$ is connected to at most one entity of $E2$ and vice versa.

- **One-to-Many** -- An entity of $E1$ may be connected to one or more entities of $E2$, but an entity of $E2$ can only be mapped to at most one entity of $E1$.

- **Many-to-Many** -- An entity of $E1$ may be linked to one or more entities of $E2$, and vice versa.
Participation Constraints

Describe how entities of $E$ participate in a relationship $R$:

- **Partial (optional) participation**
  
  An entity of $E$ does not have to be mapped to another entity through the relationship.

- **Total (mandatory) participation**
  
  Every entity of $E$ must be connected through the relationship to other entity (or entities).
Combined Notation

- **one-to-one**
  ![One-to-One Diagram]

- **one-to-many**
  ![One-to-Many Diagram]

- **many-to-many**
  ![Many-to-Many Diagram]
Relationship Attribute

// Describes the association

- A adverb or noun in requirement specification

```
Instructor ─── teaches ─── Course
    sect#       time
    room#       date
```
N-ary Relationships

- A link must associate with all participants
- Cardinality is with respect to individual relationships
- A N-ary relationship is not equivalent to N binary relationships
Weak Entity Types

- Can't exist in DB independently
- Must be identified by its owner
  - Owner entity type
  - Identify relationship
  - Partial key
  - Total participation

Diagram:

- Student is related to Car via Registers
- Student participates in Registers with 1 relationship
- Car participates in Registers with N relationship
- Attributes: Parking Lot#, Date, Color, PNum
Other E-R Models

Entity 1 -- Relationship -- Entity 2

attr. 1 attr. 2

Entity 1 -- Relationship -- Entity 2

Entity 1 \( (1:m) \) -- Relationship \( (0:5) \) -- Entity 2
Pros and Cons of E-R

Advantages

- Simple and easy to understand.
- Very popular.
- Semantic richer than classical data models.

Disadvantages:

- Not a formally defined data model.
- Deals with some integrity constraints.
- Difficult to distinguish entities from relationships.
- Has redundant modeling information.
Is-A Relationship

Generalization and specialization hierarchy

- Supertypes -- Hide the differences of subtypes
- Subtypes -- Reveal specific properties

Generalization

Person

Student

UnderGrad

Grad

Staff

Faculty

Specialization

Secretary
Inheritance

- A supertype contains the common properties of all its subtype entities.
- Subtype inherits properties of its supertype and may have its own properties.

Diagram:

- Person
  - Employee
    - Project Manager
    - Project Member
  - Dependent
    - Birth-date

Properties:
- Person: ID, Name, Address
- Employee: Salary, Skill
- Project Manager: Ratio-of-success, ...
- Project Member: Languages, ...
- Dependent: Birth-date
The American Airlines Company publishes a monthly flight log report that tracks which type of aircraft and the number of hours that were flown by an individual pilot. A separate report is prepared for each pilot and is used to monitor pilot flight proficiency for the two types of aircraft (fixed-wing and rotorcraft) which a pilot may be qualified to fly.

The following business rules apply to this report. Pilots may be assigned to fly different aircraft each day by the flight scheduling manager. Each aircraft has a single crew chief permanently assigned to perform maintenance on the aircraft, although a crew chief may crew more than one aircraft. Each aircraft is identified by an aircraft number. There are several aircraft types in the fleet. An aircraft number is unique within an aircraft type.

Identify the entities in this situation and draw an E-R diagram of the entities to include their relationships and any attributes identified in this example. Give examples of additional attributes that might be associated with each entity.